

A Parameterized Runtime Analysis of Evolutionary Algorithms for the Euclidean Traveling Salesperson Problem

Andrew M. Sutton and Frank Neumann

School of Computer Science, University of Adelaide
Adelaide, SA 5005, Australia

Abstract

Parameterized runtime analysis seeks to understand the influence of problem structure on algorithmic runtime. In this paper, we contribute to the theoretical understanding of evolutionary algorithms and carry out a parameterized analysis of evolutionary algorithms for the Euclidean traveling salesperson problem (Euclidean TSP).

We investigate the structural properties in TSP instances that influence the optimization process of evolutionary algorithms and use this information to bound the runtime of simple evolutionary algorithms. Our analysis studies the runtime in dependence of the number of inner points k and shows that $(\mu + \lambda)$ evolutionary algorithms solve the Euclidean TSP in expected time $O((\mu/\lambda) \cdot n^3 \gamma(\epsilon) + n \gamma(\epsilon) + (\mu/\lambda) \cdot n^{4k} (2k - 1)!)$ where γ is a function of the minimum angle ϵ between any three points.

Finally, our analysis provides insights into designing a mutation operator that improves the upper bound on expected runtime. We show that a mixed mutation strategy that incorporates both 2-opt moves and permutation jumps results in an upper bound of $O((\mu/\lambda) \cdot n^3 \gamma(\epsilon) + n \gamma(\epsilon) + (\mu/\lambda) \cdot n^{2k} (k - 1)!)$ for the $(\mu + \lambda)$ EA.

1 Introduction

In many real applications, the inputs of an NP-hard combinatorial optimization problem may be structured or restricted in such a way that it becomes tractable to solve in practice despite having a worst-case exponential time bound. Parameterized analysis seeks to address this by expressing algorithmic runtime in terms of an additional hardness parameter that isolates the source of exponential complexity in the problem structure. In this paper, we study the application of evolutionary algorithms (EAs) to the Euclidean Traveling Salesperson Problem (TSP) and consider the runtime of such algorithms as a function of both problem size and a further parameter that influences how hard the problem is to solve by an EA.

1.1 The Euclidean traveling salesperson problem

Iterative heuristic methods (such as local search and evolutionary algorithms) that rely on the exchange of a few edges such as the well-known 2-opt (or 2-change) operator are popular choices for solving large scale TSP instances in practice. This is partly due to the fact that they have a simple implementation and typically perform well empirically. However, for these algorithms, theoretical understanding still remains limited. Worst-case analyses demonstrate the existence of

instances on which the procedure can be inefficient. Chandra, Karloff, and Tovey [3], building from unpublished results due to Lueker, have shown that local search algorithms employing a k -change neighborhood operator can take exponential time to find a locally optimal solution. Even in the Euclidean case, Englert, Röglin, and Vöcking [11] have recently shown that a local search algorithm employing inversions can take worst-case exponential time to find tours which are locally optimum.

If the search operator is restricted to specialized 2-opt moves that remove only edges that intersect in the plane, van Leeuwen and Schoone [30] proved that a tour that has no such planar intersections can be reached in $O(n^3)$ moves, even if a move introduces further intersecting edges. Since determining which edges are intersecting can take quadratic time, a locally optimal tour can be found in time $O(n^5)$. We point out that a local optimum in this restricted neighborhood does not necessarily correspond to a local optimum in the general 2-opt neighborhood.

If the vertices are distributed uniformly at random in the unit square, Chandra, Karloff, and Tovey [3] showed that the expected time to find a locally optimal solution is bounded by $O(n^{10} \log n)$. More generally, for so-called ϕ -perturbed Euclidean instances, Englert, Röglin, and Vöcking [11] proved that the expected time to find a locally optimum solution is bounded by $O(n^{4+1/3} \log(n\phi)\phi^{8/3})$. These results also imply similar bounds for simple ant colony optimization algorithms as shown in [16].

To allow for a deeper insight into the relationship between problem instance structure and algorithmic runtime, we appeal in this paper to the theory of parameterized complexity [7]. Rather than expressing the runtime solely as a function of problem size, parameterized analysis decomposes the runtime into further parameters that are related to the structure of instances. The idea is to find parameters that partition off the combinatorial explosion that leads to exponential runtimes [8].

In the context of TSP, a number of parameterized results currently exist. Deineko et al. [5] showed that, if a Euclidean TSP instance with n vertices has k vertices interior to the convex hull, there is a dynamic programming algorithm that can solve the instance in time bounded by $g(k) \cdot n^{O(1)}$ where g is a function that depends only on k . This means that this parameterization belongs to the complexity class FPT, the class of parameterized problems that are considered fixed-parameter tractable. Of course, membership in FPT depends strongly on the parameterization itself. For example, the problem of searching the k -change neighborhood for metric TSP is hard for W[1] due to Marx [19]. Therefore, the latter parameterization is not likely to belong to FPT.

1.2 Computational complexity of evolutionary algorithms

Initial studies on the computational complexity of evolutionary algorithms consider their runtime on classes of artificial pseudo-Boolean functions [10, 13, 14, 32]. The goal of these studies is to consider the impact of the different modules of an evolutionary algorithm and to develop new methods for their analysis. This early work was instrumental in establishing a rigorous understanding of the behavior of evolutionary algorithms on simple functions, for identifying some classes of problems that simple EAs can provably solve in expected polynomial time [10], and for disproving widely accepted conjectures (e.g., that evolutionary algorithms are always efficient on unimodal functions [9]).

More recently, classical polynomial-time problems from combinatorial optimization such as minimum spanning trees [22, 20] and shortest paths [25, 6, 2] have been considered. In this case, one does not hope to beat the best problem-specific algorithms for classical polynomial solvable problems. Instead, these studies provide interesting insights into the search behavior of these algorithms

and show that many classical problems are solved by general-purpose algorithms such as evolutionary algorithms in expected polynomial time.

Research on NP-hard combinatorial optimization problems such as makespan scheduling, covering problems, and multi-objective minimum spanning trees [21, 31] show that evolutionary algorithms can achieve good approximations for these problems in expected polynomial time. In the case of the TSP, Theile [29] has proved that a $(\mu + 1)$ EA based on dynamic programming can exactly solve the TSP in at most $O(n^3 2^n)$ steps when μ is allowed to be exponential in n . For a comprehensive presentation of the different results that have been achieved see, e.g., the recent text of Neumann and Witt [23].

Algorithmic runtime on NP-hard problems can be studied in much sharper detail from the perspective of parameterized analysis, and this has only recently been started in theoretical work on evolutionary algorithms. Parameterized results have been obtained for the vertex cover problem [18], the problem of computing a spanning tree with a maximal number of leaves [17], variants of maximum 2-satisfiability [26], and makespan scheduling [28].

1.3 Our results

In this paper, we carry out a parameterized complexity analysis for evolutionary algorithms for the Euclidean TSP. We prove upper bounds on the expected runtime of two classical EAs based on 2-opt mutation in the context of the TSP parameterization of Deĭneko et al. [5], that is, as a function of the number of points that lie on the interior of the convex hull. Our results are for the $(\mu + \lambda)$ EA which operates on a population of μ permutations (candidate Hamiltonian cycles) and produces λ offspring in each generation using a mutation operator based on 2-opt. This analysis provides further insights into the optimization process that allows us to design a mixed mutation operator that uses both 2-opt moves and permutation jumps and improves the upper bound on expected runtime of the $(\mu + \lambda)$ EA.

By setting $\mu = \lambda = 1$ and changing the mutation operator to single random 2-opt moves, we also prove parameterized runtime bounds for randomized local search (RLS): a randomized hill-climber on the space of permutations. In this case, we present results for the expected time for RLS to converge to a locally optimal tour in terms of 2-opt moves. Specialized results for RLS and the $(1 + 1)$ EA using 2-opt mutation appear in a conference version of this paper presented at AAAI 2012 [27].

The paper is organized as follows. In Section 2 we introduce the problem, algorithm and analysis. In Section 3 we study structural properties of the Euclidean TSP. In Section 4 we study the runtime of the $(\mu + \lambda)$ EA and RLS on Euclidean TSP instances whose points lie in convex position, i.e., have no inner points. In Section 5 we then prove rigorous runtime bounds for the algorithms as a function of the number of inner points in an instance. We conclude the paper in Section 6.

2 Preliminaries

Let V be a set of n points in the plane labeled as $[n] = \{1, \dots, n\}$ such that no three points are collinear. We consider the complete, weighted Euclidean graph $G = (V, E)$ where E is the set of all 2-sets from V . The weight of an edge $\{u, v\} \in E$ is equal to $d(u, v)$: the Euclidean distance separating the points. The goal is to find a set of n edges of minimum weight that form a

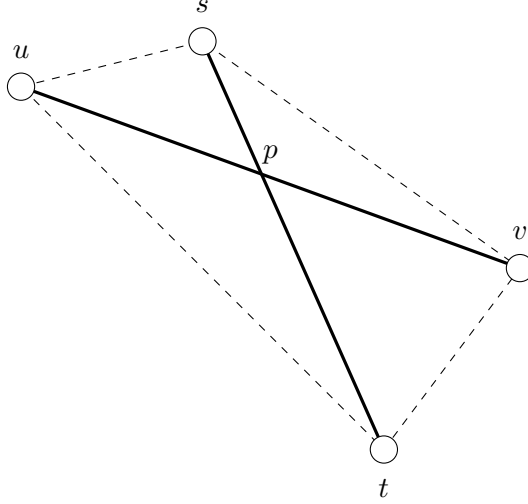


Figure 1: Intersecting edges form the diagonals of a convex quadrilateral.

Hamiltonian cycle in G . A candidate solution of the TSP is a permutation x of V which we consider as a sequence of distinct elements $x = (x_1, x_2, \dots, x_n)$, such that $x_i \in [n]$. The Hamiltonian cycle in G induced by such a permutation is the set of n edges

$$C(x) = \{\{x_1, x_2\}, \{x_2, x_3\}, \dots, \{x_{n-1}, x_n\}, \{x_n, x_1\}\}.$$

The optimization problem is to find a permutation x which minimizes the fitness function

$$f(x) = \sum_{\{u,v\} \in C(x)} d(u,v). \quad (1)$$

Geometrically, it will often be convenient to consider an edge $\{u, v\}$ as the unique planar line segment with end points u and v . We say a pair of edges $\{u, v\}$ and $\{s, t\}$ *intersect* if they cross at a point in the Euclidean plane. An important observation, which we state here without proof, is that any pair of intersecting edges form the diagonals of a convex quadrilateral in the plane (see Figure 1).

Proposition 1. *If $\{u, v\}$ and $\{s, t\}$ intersect at a point p , they form the diagonals of a convex quadrilateral described by points u, s, v , and t . Hence edges $\{s, u\}$, $\{s, v\}$, $\{t, v\}$ and $\{t, u\}$ form a set of edges that mutually do not intersect.*

Definition 1. *A tour $C(x)$ is called intersection-free if it contains no pairs of edges that intersect.*

2.1 Parameterized analysis

Parameterized complexity theory is an extension to traditional computational complexity theory in which the analysis of hard algorithmic problems is decomposed into parameters of the problem input. This approach illuminates the relationship between hardness and different aspects of problem structure because it often isolates the source of exponential complexity in NP-hard problems.

A parameterization of a problem is a mapping of problem instances into the set of natural numbers. We are interested in expressing algorithmic complexity in terms of both problem size and the extra parameter. Formally, let L be a language over a finite alphabet Σ . A *parameterization* of L is a mapping $\kappa : \Sigma^* \rightarrow \mathbb{N}$. The corresponding *parameterized problem* is the pair (L, κ) .

For a string $x \in \Sigma^*$, let $k = \kappa(x)$ and $n = |x|$. An algorithm deciding $x \in L$ in time bounded by $g(k) \cdot n^{O(1)}$ is called a *fixed-parameter tractable* (or fpt-) algorithm for the parameterization κ . Here $g : \mathbb{N} \rightarrow \mathbb{N}$ is an arbitrary but computable function. Similarly, an algorithm that decides a parameterized problem (L, κ) in time bounded by $n^{g(k)}$ is called an XP-algorithm.

When working with the runtime of randomized algorithms such as evolutionary algorithms, one is often interested in the random variable T which somehow measures the number of steps the algorithm must take to decide a parameterized problem. A randomized algorithm with *expected* optimization time $E(T) \leq g(k) \cdot n^{O(1)}$ (respectively, $E(T) \leq n^{g(k)}$) is a randomized fpt-algorithm (respectively, XP-algorithm) for the corresponding parameterization κ .

In the case of the Euclidean TSP on a set of points V , we want to express the runtime complexity as a function of n and k where $n = |V|$ and k is the number of vertices that lie in the interior of the convex hull of V . We will hereafter refer to these points as the *inner points* of V . For the corresponding optimization problem, we are interested in the runtime until the optimal solution is located.

2.2 Simple evolutionary algorithms

Randomized search heuristics such as randomized local search and evolutionary algorithms that are tasked to solve TSP instances are usually designed to iteratively search the space of permutations in order to minimize the fitness function defined in Equation (1). Each permutation corresponds to a particular Hamiltonian cycle in the graph. To move through the space of candidate solutions, move or mutation operators are often constructed based on some kind of elementary operations on the set of permutations on $[n]$. In this paper, we will consider two such operations: *inversions* and *jumps* which we define as follows.

Definition 2. The inversion operation σ_{ij}^I transforms permutations into one another by segment reversal. A permutation x is transformed into a permutation $\sigma_{ij}^I[x]$ by inverting the subsequence in x from position i to position j where $1 \leq i < j \leq n$.

$$\begin{aligned} x &= (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n) \\ \sigma_{ij}^I[x] &= (x_1, \dots, x_{i-1}, x_j, x_{j-1}, \dots, x_{i+1}, x_i, x_{j+1}, \dots, x_n) \end{aligned}$$

In the space of Hamiltonian cycles, the permutation inversion operation is essentially identical to the well-known 2-change (or 2-opt) operation for TSP. The usual effect of the inversion operation is to delete the two edges $\{x_{i-1}, x_i\}$ and $\{x_j, x_{j+1}\}$ from $C(x)$ and reconnect the tour $C(\sigma_{ij}^I[x])$ using edges $\{x_{i-1}, x_j\}$ and $\{x_i, x_{j+1}\}$ (see Figure 2). Here and subsequently, we consider arithmetic on the indices to be modulo n , i.e., $1 - 1 = n$ and $n + 1 = 1$. Since the underlying graph G is undirected, when $(i, j) = (1, n)$, the operation has no effect since the current tour is only reversed. There is also no effect when $(i, j) \in \{(2, n), (1, n - 1)\}$. In this case, it is straightforward to check that the edges removed from $C(x)$ are equal to the edges replaced to create $C(\sigma_{ij}^I[x])$.

Definition 3. The jump operation σ_{ij}^J transforms permutations into one another by position shifts. A permutation x is transformed into a permutation $\sigma_{ij}^J[x]$ by moving the element in position i into

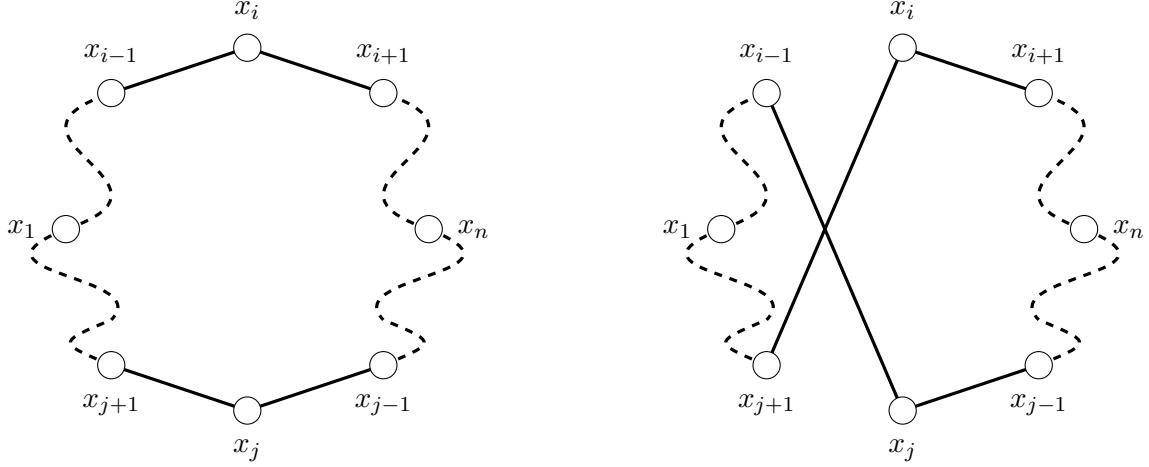


Figure 2: The effect of the inversion operation on a Hamiltonian cycle.

position j while the other elements between position i and position j are shifted in the appropriate direction. Without loss of generality, suppose $i < j$. Then,

$$\begin{aligned}
 x &= (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n) \\
 \sigma_{ij}^J[x] &= (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{j-1}, x_i, x_j, x_{j+1}, \dots, x_n) \\
 \sigma_{ji}^J[x] &= (x_1, \dots, x_{i-1}, x_j, x_i, x_{i+1}, \dots, x_{j-1}, x_{j+1}, \dots, x_n)
 \end{aligned}$$

Since $\sigma_{i(i+1)}^J$ and $\sigma_{(i+1)i}^J$ have the same effect, there are $n(n-1) - (n-1) = (n-1)^2$ unique jump operations. The *jump* operator σ_{ij}^J was used by Scharnow, Tinnefeld and Wegener [25] in the context of runtime analysis of evolutionary algorithms on permutation sorting problems.

In this paper, we consider simple mutation-only evolutionary algorithms that operate on permutations as follows.

- 1. Initialization:** Generate a set of μ permutations on $[n]$ uniformly at random.
- 2. Mutation:** Generate a set of λ “offspring” permutations by applying some type of randomized move or mutation operator based on the above operations.
- 3. Selection:** Select the fittest μ permutations out of the both parent and offspring population and continue at line 2.

The general form of such a mutation-only evolutionary algorithm is typically called a $(\mu + \lambda)$ EA. Algorithm 1 illustrates the $(\mu + \lambda)$ EA. Note that here, the **select** function selects the μ best (with respect to fitness) permutations from the parent population P and the offspring population P' . This selection mechanism ensures that best-so-far individual found by generation t remains in the population at generation t , i.e., the $(\mu + \lambda)$ EA exhibits *elitism*.

To generate offspring using the previously introduced inversion operation, the $(\mu + \lambda)$ RA employs a mutation operator that applies a number of random 2-opt moves that is drawn from a Poisson distribution. This mutation operator, called **2-opt-mutation**, is outlined in Function 2.

If we restrict the mutation operation to a single inversion move and set $\mu = \lambda = 1$, the resulting algorithm is *randomized local search* (RLS), which is simply a randomized hill-climber in

Algorithm 1: The $(\mu + \lambda)$ EA.

```
1 Choose a multiset  $P$  of  $\mu$  random permutations on  $[n]$ ;  
2 repeat forever  
3    $P' \leftarrow \{\}$ ;  
4   repeat  $\lambda$  times  
5     choose  $x$  uniformly at random from  $P$ ;  
6      $y \leftarrow \text{mutate}(x)$ ;  
7      $P' \leftarrow P' \uplus \{y\}$ ;  
8    $P \leftarrow \text{select}(P \uplus P')$  ;
```

Function 2: 2-opt-mutation(x)

```
input : A permutation  $x$   
output: A permutation  $y$   
  
1  $y \leftarrow x$ ;  
2 draw  $s$  from a Poisson distribution with parameter 1;  
3 perform  $s + 1$  random inversion operations on  $y$ ;  
4 return  $y$ ;
```

the space of permutations using 2-opt moves. RLS, illustrated Algorithm 3, operates by iteratively applying random inversion operations to a permutation in order to try and improve the fitness of the corresponding tours. Unlike the $(\mu + \lambda)$ EA, RLS can only generate immediate inversion neighbors so it can become trapped in local optima.

Algorithm 3: Randomized Local Search (RLS).

```
1 Choose a random permutation  $x$  on  $[n]$ ;  
2 repeat forever  
3   choose a random distinct pair of elements  $(i, j)$  from  $[n]$ ;  
4    $y \leftarrow \sigma_{ij}^I[x]$ ;  
5   if  $f(y) \leq f(x)$  then  $x \leftarrow y$ 
```

2.3 Runtime analysis

Evolutionary algorithms are simply computational methods that rely on random decisions so we consider them here as special cases of *randomized algorithms*. To analyze the running time of such an algorithm, we examine the sequence of best-so-far solutions it discovers during execution

$$(x^{(1)}, x^{(2)}, \dots, x^{(t)}, \dots)$$

as an infinite stochastic process where $x^{(t)}$ denotes the best permutation (in terms of fitness) in the population at iteration t . The goal of runtime analysis is to study the random variable that equals the first time t when $x^{(t)}$ is a candidate solution of interest (for example, an optimal solution).

The *optimization time* of a randomized algorithm is a random variable

$$T = \inf\{t \in \mathbb{N} : f(x^{(t)}) \text{ is optimal}\}. \quad (2)$$

In the case of the $(\mu + \lambda)$ EA, this corresponds to the number of generations (iterations of the mutation, evaluation, selection process) that occur before an optimal solution has been introduced to the population. This is somewhat distinct the traditional measure of the number of explicit calls to the fitness function [23, 1]. However, in the case of the $(\mu + \lambda)$ EA, this metric can be obtained from T by $T_f = \mu + \lambda T$, since we need μ fitness function calls to evaluate the initial population and each generation requires evaluating an additional λ individuals. We discuss this further in section 5.

In this paper, we will estimate the *expected optimization time* of the $(\mu + \lambda)$ EA. This quantity is calculated as $E(T)$, the expectation of T . Since RLS can become trapped in local optima, its expected optimization time is not necessarily finite. In this case, we introduce the random variable

$$T_{loc} = \inf\{t \in \mathbb{N} : x^{(t)} \text{ has no improving 2-opt neighbors}\}. \quad (3)$$

and estimate the expected time RLS takes to reach a locally optimal solution.

Definition 4. Let α be an indicator function defined on permutations of $[n]$ as

$$\alpha(x) = \begin{cases} 1 & \text{if } C(x) \text{ contains intersections;} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 5. Let β be an indicator function defined on permutations of $[n]$ as

$$\beta(x) = \begin{cases} 1 - \alpha(x) & \text{if } f(x) \text{ non-optimal} \\ 0 & \text{otherwise.} \end{cases}$$

The random variable corresponding to optimization time can be expressed as the infinite series

$$T = \sum_{t=1}^{\infty} (\alpha(x^{(t)}) + \beta(x^{(t)})). \quad (4)$$

In order to characterize the behavior of evolutionary algorithms and express their expected runtime in terms of the number of points n and the number of inner points k , we analyze the Markov chain generated by the algorithm. We construct the Markov chain as follows. Given a point set V , Each permutation x on $[n]$ that corresponds to a tour $C(x)$ that is non-optimal is a unique state in the Markov chain. Finally, every permutation that corresponds to an optimal tour in V is associated with a single absorbing state *opt*. We then bipartition the state space (minus *opt*) into two sets S_α and S_β where

$$S_\alpha = \{x : C(x) \text{ contains intersecting edges}\},$$

and

$$S_\beta = \{x : C(x) \text{ is intersection-free}\} \setminus \{\text{opt}\},$$

In terms of the Markov chain, the optimization time T is the first hitting time of the state *opt*. We will need the following two preparatory lemmas, the first of which is analogous to the additive drift result due to He and Yao [13].

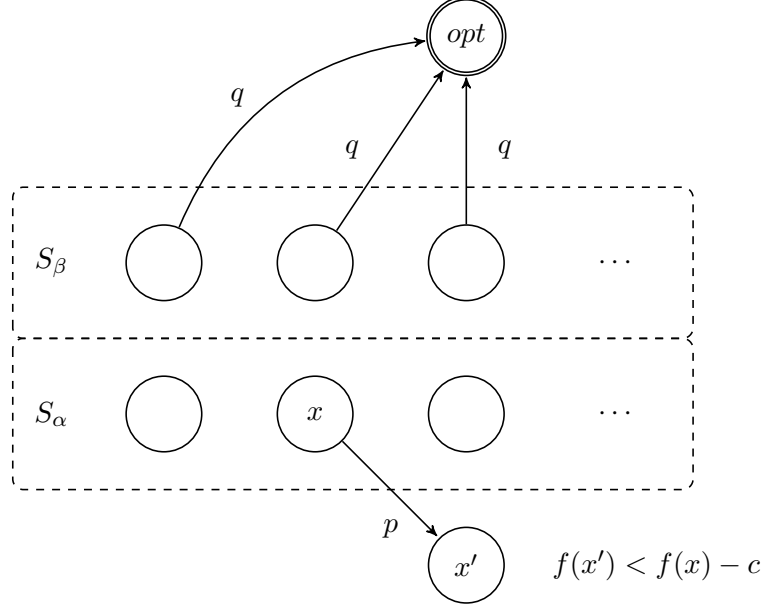


Figure 3: Partitioned Markov chain with transition probability bounds. Note that x' can be in either S_α or S_β .

Lemma 1. *If there are constants $0 < p < 1$ and $c > 0$ such that for any $x \in S_\alpha$, the transition probability from x to some x' with $f(x') < f(x) - c$ is bounded below by p , then*

$$\mathbb{E} \left(\sum_{t=1}^{\infty} \alpha(x^{(t)}) \right) \leq p^{-1} n (d_{\max} - d_{\min}) / c,$$

where d_{\max} and d_{\min} are the maximum and minimum distances between any two points in V , respectively.

Proof. By hypothesis, since $p > 0$ for any $x \in S_\alpha$ there is a nonzero probability the algorithm exits the state x and never returns since the sequence of best-so-far solutions increases monotonically in fitness. Therefore, the expectation in the claim exists and is finite.

Consider the finite stochastic process $(y^{(1)}, y^{(2)}, \dots, y^{(m)})$, which is defined as the restriction of $(x^{(1)}, x^{(2)}, \dots)$ constructed by taking only permutations $y^{(t)} \in S_\alpha$ in the same order. It follows that

$$\mathbb{E} \left(\sum_{t=1}^{\infty} \alpha(x^{(t)}) \right) = \mathbb{E}(m).$$

Let $y^{(1)}$ be the first state in the restricted stochastic process. Since every state has a transition to a state that improves the fitness by at least c with transition probability at least p , the expected number of states $y^{(t)}$ with $f(y^{(1)}) \geq f(y^{(t)}) \geq f(y^{(1)}) - c$ is bounded by p^{-1} . Continuing this argument, the expected waiting time until the fitness of any state $y^{(t)}$ is improved by at least c is at most p^{-1} .

For any arbitrary permutation x , $nd_{\max} \geq f(x) \geq nd_{\min}$ so there can be at most $n(d_{\max} - d_{\min})/c$ such improvements possible. \square

Lemma 2. *If there is a constant $0 < q < 1$ such that for any $x \in S_\beta$, the transition probability from x to opt is bounded below by q , then*

$$\mathbb{E} \left(\sum_{t=1}^{\infty} \beta(x^{(t)}) \right) \leq q^{-1}$$

Proof. Again, since $q > 0$ for any $x \in S_\beta$ there is a nonzero probability the algorithm exits the state x and transits to the absorbing state opt . It follows that the expected time spent in states contained in the S_β partition is bounded above by q^{-1} . \square

In the next section, we will carefully analyze properties of the Euclidean TSP to find suitable values for p , q , and c . This will allow us to bound the expected runtime in terms of n and k .

3 Structural properties

We now examine some useful structural properties of Euclidean TSP instances related to the inversion operator. We also introduce some structural constraints that will later facilitate the parameterized analysis. We begin by pointing out that if a tour is not intersection-free, an intersection can always be removed by an inversion. This notion is captured by the following lemma.

Lemma 3. *Let x be a permutation such that $C(x)$ is not intersection-free. Then there exists an inversion that removes a pair of intersecting edges and replaces them with a pair of non-intersecting edges.*

Proof. Suppose $\{x_{i-1}, x_i\}$ and $\{x_j, x_{j+1}\}$ intersect in $C(x)$. Let $y = \sigma_{ij}^I[x]$. Then

$$\begin{aligned} C(x) \setminus C(y) &= \{\{x_{i-1}, x_i\}, \{x_j, x_{j+1}\}\}, & \text{and} \\ C(y) \setminus C(x) &= \{\{x_{i-1}, x_j\}, \{x_i, x_{j+1}\}\}. \end{aligned}$$

By Proposition 1, since $\{x_{i-1}, x_i\}$ and $\{x_j, x_{j+1}\}$ intersect, the two new edges introduced to $C(y)$ by σ_{ij}^I do not intersect. Note that it is still possible that the introduced edges intersect with some of the remaining edges in $C(y)$. \square

We denote by $\mathfrak{H}(V) \subseteq V$ the convex hull of V . A permutation x respects hull-order if any two points in the subsequence of x induced by $\mathfrak{H}(V)$ are consecutive in x if and only if they are consecutive on the hull.

Lemma 4. *If $C(x)$ is intersection-free, then x respects hull-order.*

Proof. This follows immediately from the proof of Theorem 2 in [24]. \square

Lemma 4 entails the following bound on the cardinality of S_β , the set of permutations that correspond to intersection-free tours.

Lemma 5. *Suppose $|V \setminus \mathfrak{H}(V)| = k$. Then $|S_\beta| \leq (n - k)^k k!$.*

Proof. For any set of $1 < p < n$ points, there are $p^{n-p}(n-p)!$ permutations in which the p points remain in the same order. Since $|\mathfrak{H}(V)| = (n - k)$, there are exactly $(n - k)^k k!$ permutations that respect hull-order. Since each intersection-free tour must respect hull-order, we have the claimed bound. \square

We also can derive from Lemma 4 the following convenient bound on the minimal number of operations necessary to transform an intersection-free tour into a permutation that corresponds to a globally optimal tour.

Lemma 6. *Suppose $|V \setminus \mathfrak{H}(V)| = k$ and $C(x)$ is an intersection-free tour on V . Then there is a sequence of at most k jump operations that transforms x into an optimal permutation.*

Proof. By Lemma 4, since $C(x)$ is intersection-free, it must be hull-respecting. Let x^* be an optimal permutation such that the elements in $\mathfrak{H}(V)$ have the same linear order in x^* as they do in x . Then x can be transformed into x^* by moving each of the k inner points into their correct position. \square

Lemma 7. *Suppose $|V \setminus \mathfrak{H}(V)| = k$ and $C(x)$ is an intersection-free tour on V . Then there is a sequence of at most $2k$ inversions that transforms x into an optimal permutation.*

Proof. By Lemma 6, k jump operations are enough to transform x into an optimal permutation. The claim then immediately follows from the fact that any jump operation can be simulated by at most two consecutive inversion operations. In particular,

$$\sigma_{ij}^J[x] = \begin{cases} \sigma_{ij}^I[x] & \text{if } |i - j| = 1; \\ \sigma_{i(j-1)}^I[\sigma_{ij}^I[x]] & \text{if } i - j < 1; \\ \sigma_{(j+1)i}^I[\sigma_{ji}^I[x]] & \text{if } i - j > 1. \end{cases}$$

Since a sequence of at most k jump operations can be simulated by a sequence of at most $2k$ inversion operations, the claim is proved. \square

A challenge to the the runtime analysis of algorithms that employ edge exchange operations such as 2-opt is that, when points are allowed in arbitrary positions, the minimum change in fitness between neighboring solutions can be made arbitrarily small. Indeed, proof techniques for worst-case analysis often leverage this fact [11]. To circumvent this, we impose bounds on the angles between points, which allows us to express runtime results as a function of trigonometric expressions involving these bounds. Momentarily, we will refine this further by introducing a class of TSP instances embedded in an $m \times m$ grid. In that case, we will see that the resulting trigonometric expression is bounded by a polynomial in m .

We say V is *angle-bounded* by $\epsilon > 0$ if for any three points $u, v, w \in V$, $0 < \epsilon < \theta < \pi - \epsilon$ where θ denotes the angle formed by the line from u to v and the line from v to w . This allows us to express a bound in terms of ϵ on the change in fitness from a move that removes an inversion.

Lemma 8. *Suppose V is angle-bounded by ϵ . Let x be a permutation such that $C(x)$ is not intersection-free. Let $y = \sigma_{ij}^I[x]$ be the permutation constructed from an inversion on x that replaces two intersecting edges in $C(x)$ with two non-intersecting edges.¹ Then, if d_{\min} denotes the minimum distance between any two points in V , $f(x) - f(y) > 2d_{\min} \left(\frac{1 - \cos(\epsilon)}{\cos(\epsilon)} \right)$.*

Proof. The inversion σ_{ij}^I removes intersecting edges $\{u, v\}$ and $\{s, t\}$ from $C(x)$ and replaces them with the pair $\{s, u\}$ and $\{t, v\}$ to form $C(y)$. We label the point at which the original edges intersect as p .

¹Lemma 3 guarantees the existence of such an inversion.

Denote as θ_u and θ_v the angles between the line segments that join at each point u and v , respectively. Since all angles are strictly positive, the points u , s , and p form a nondegenerate triangle with angles θ_s , θ_u , and $(\pi - (\theta_s + \theta_u))$. By the law of sines we have

$$\frac{d(s, u)}{\sin(\pi - (\theta_s + \theta_u))} = \frac{d(s, u)}{\sin(\theta_s + \theta_u)} = \frac{d(u, p)}{\sin(\theta_s)} = \frac{d(s, p)}{\sin(\theta_u)}.$$

Hence,

$$d(u, p) + d(s, p) = d(s, u) \left(\frac{\sin(\theta_s) + \sin(\theta_u)}{\sin(\theta_s + \theta_u)} \right). \quad (5)$$

Since u , s , and p form a triangle, $0 < (\theta_s + \theta_u) < \pi$ and we have

$$\begin{aligned} 0 < \sin(\theta_s) < 1 & \quad \text{since } 0 < \theta_s < \pi, \\ 0 < \sin(\theta_u) < 1 & \quad \text{since } 0 < \theta_u < \pi, \\ 0 < \sin(\theta_s + \theta_u) < 1 & \quad \text{since } 0 < \theta_s + \theta_u < \pi. \end{aligned}$$

Furthermore, since V is angle-bounded by $0 < \epsilon < \pi - \epsilon$, by (5),

$$d(u, p) + d(s, p) > d(s, u) \left(\frac{\sin(\epsilon) + \sin(\epsilon)}{\sin(\epsilon + \epsilon)} \right) > d(s, u). \quad (6)$$

Since there is also a nondegenerate triangle formed by the points t , v , and p , a symmetric argument holds and thus

$$d(t, p) + d(v, p) > d(t, v) \left(\frac{\sin(\epsilon) + \sin(\epsilon)}{\sin(\epsilon + \epsilon)} \right) > d(t, v). \quad (7)$$

Combining Equations (6) and (7) we have

$$\begin{aligned} f(x) - f(y) &= [d(u, v) + d(s, t)] - [d(t, v) + d(s, u)] \\ &= d(u, p) + d(v, p) + d(t, p) + d(s, p) - [d(t, v) + d(s, u)] \\ &> [d(t, v) + d(s, u)] \left(\frac{2 \sin(\epsilon)}{\sin(2\epsilon)} \right) - [d(t, v) + d(s, u)] > 0 \end{aligned}$$

The constraint that the difference is strictly positive follows directly from Equations (6) and (7). Hence,

$$\begin{aligned} f(x) - f(y) &> [d(t, v) + d(s, u)] \left(\frac{2 \sin(\epsilon)}{\sin(2\epsilon)} - 1 \right) \\ &\geq 2d_{\min} \left(\frac{2 \sin(\epsilon)}{\sin(2\epsilon)} - 1 \right) = 2d_{\min} \left(\frac{1 - \cos(\epsilon)}{\cos(\epsilon)} \right). \quad \square \end{aligned}$$

The lower bound on the angle between any three points in V provides a constraint on how small the change in fitness between neighboring inversions can be. This lower bound is useful in the case of a *quantized* point-set. That is, when the points can be embedded on an $m \times m$ grid as illustrated in Figure 4.

Quantization, for example, occurs when the x and y coordinates of each point in the set are rounded to the nearest value in a set of m equidistant values (e.g., integers). We point out that it is still important that the quantization preserves the constraint on collinearity since collinear points violate a nonzero angle bound. We have the following lemma.

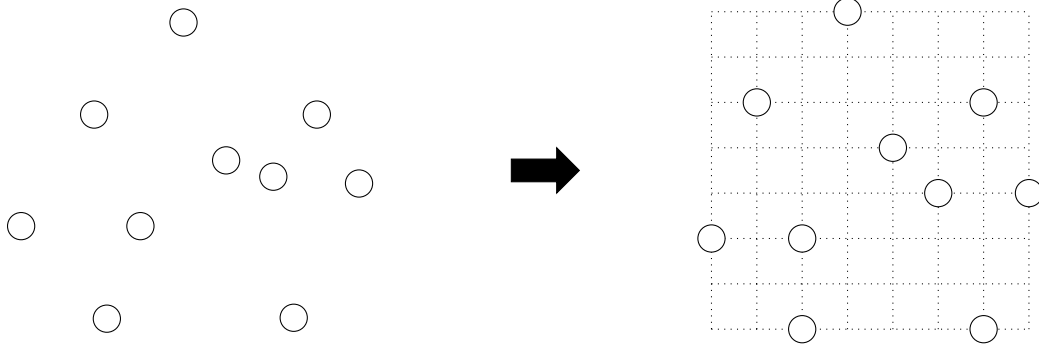


Figure 4: Quantizing a set of n points in the plane onto an $m \times m$ grid.

Lemma 9. *Suppose V is a set of points that lie on an $m \times m$ unit grid, no three collinear. Then V is angle-bounded by $\arctan(1/(2(m-2)^2))$.*

Proof. The grid imposes a coordinate system on V in which the concept of line slope is well-defined. Let $u, v, w \in V$ be arbitrary points. We consider the angle θ at point v formed by the lines from v to u and v to w . Let s_1 and s_2 denote the slope of these lines, respectively. If the slopes are of opposite sign, then $\theta \geq 2 \arctan((m-1)^{-1})$ since the lines form hypotenuses of two right triangles with adjacent sides of length at most $m-1$ and opposite sides with length at least 1 (see Figure 5).

We now consider the case where the slopes are nonnegative. The nonpositive case is handled identically (or by simply changing the sign of the slopes by the appropriate transformation). Without loss of generality, assume $s_1 > s_2 \geq 0$. Equality is impossible since u, v , and w cannot be collinear. Since the points lie on an $m \times m$ grid, s_1 and s_2 must be ratios of whole numbers at most $m-1$, say $s_1 = a/b$ and $s_2 = c/d$. The angle at point v is $\theta = \arctan(a/b) - \arctan(c/d) = \arctan\left(\frac{ad-cb}{bd+ac}\right)$. The minimum positive value for the expression $(ad-cb)/(bd+ac)$ over the integers from 0 to $m-1$ is $\frac{1}{2(m-2)^2}$. Since the inverse of the tangent is monotone, the minimum nonzero angle must be $\theta \geq \arctan(1/(2(m-2)^2))$. \square

Lemma 9 allows us to translate the somewhat awkward trigonometric expression in the claim of Lemma 8 (and subsequent lemmas that depend on it) into a convenient polynomial that can be expressed in terms of m .

Lemma 10. *Let V be a set of n points that lie on an $m \times m$ unit grid, no three collinear. Then, V is angle-bounded by ϵ where $\cos(\epsilon)/(1 - \cos(\epsilon)) = O(m^4)$.*

Proof. It follows from Lemma 9 that the angle bound on V is $\epsilon = \arctan(1/(2(m-2)^2))$. Since $\cos(\arctan(x)) = 1/\sqrt{1+x^2}$ we have

$$\frac{\cos(\epsilon)}{1 - \cos(\epsilon)} = \frac{2(m-2)^2}{\sqrt{1 + 4(m-2)^4} - 2(m-2)^2}.$$

and since $z/(\sqrt{1+z^2} - z) = O(z^2)$, setting $z = 2(m-2)^2$ completes the proof. \square

We are now ready to prove the following technical lemma for 2-opt mutation defined in Function 2. This lemma will be instrumental in proving runtime bounds later in the paper.

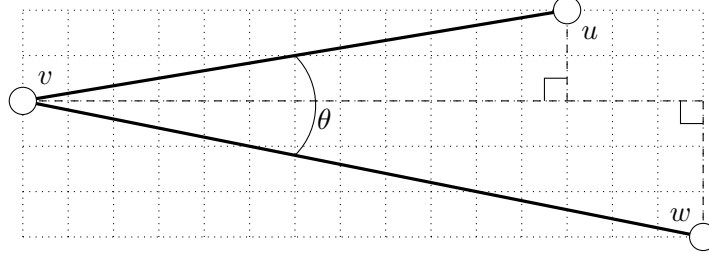


Figure 5: If the slope of the lines from v to u and v to w are of opposite sign, they form the hypotenuses of two right triangles and $\theta \geq 2 \arctan((m-1)^{-1})$.

Lemma 11. *Let V be a set of planar points in convex position angle-bounded by ϵ . We have the following.*

- (1) *For any $x \in S_\alpha$, the probability that 2-opt mutation creates an offspring y with $f(y) < f(x) - 2d_{\min}(1 - \cos(\epsilon)) / (\cos(\epsilon))$ is at least $2/(en(n-1))$.*
- (2) *For any $x \in S_\beta$, the probability that 2-opt mutation creates an optimal solution is at least $(en^{4k}(2k-1)!)^{-1}$*

Proof. For (1), suppose $x \in S_\alpha$. By Lemma 3, there is at least one pair of intersecting edges in $C(x)$ that can be removed with a single inversion operation σ_{ij}^I . Let $y = \sigma_{ij}^I[x]$. By Lemma 8, $f(y)$ satisfies the fitness bound in the claim. It suffices to bound the probability that y is produced by **2-opt-mutation**(x).

Let E_1 denote the event that Poisson mutation performs exactly one inversion (i.e., $s = 0$ in line 2 of Function 2). Let E_2 denote the event that the pair (i, j) is specifically chosen for the inversion.

We have $\Pr\{E_1\} = e^{-1}$ from the Poisson density function and $\Pr\{E_2|E_1\} \geq (n(n-1)/2)^{-1}$. Thus the probability that 2-opt mutation creates y from x is

$$\Pr\{E_1 \cup E_2\} = \Pr\{E_1\} \cdot \Pr\{E_2|E_1\} \geq 2/(en(n-1)).$$

For (2), suppose $x \in S_\beta$. Thus, $C(x)$ is intersection-free, and it follows from Lemma 7 that there are at most $2k$ inversion moves that transform x into an optimal solution.

Let E'_1 denote the event that Poisson mutation performs exactly $2k$ inversions (i.e., $s = 2k - 1$ in line 2 of Function 2). Let E'_2 denote the event that all $2k$ inversions are the correct moves that transform x into an optimal solution.

Again, from the Poisson density function, $\Pr\{E'_1\} = (e(2k-1)!)^{-1}$. Since $\Pr\{E'_2|E'_1\} \geq (n(n-1)/2)^{-2k} \geq n^{-4k}$, the probability of transforming x into an optimal solution is at least

$$\Pr\{E'_1 \cap E'_2\} = \Pr\{E'_1\} \cdot \Pr\{E'_2|E'_1\} \geq (en^{4k}(2k-1)!)^{-1}.$$

□

Finally, since the time bounds in the remainder of this paper are expressed as a function of the angle bound ϵ , and the bounds on point distance, it will be useful to define the following function.

Definition 6. Let V be a set of points angle-bounded by ϵ . We define

$$\gamma(\epsilon) = \left(\frac{d_{\max}}{d_{\min}} - 1 \right) \left(\frac{\cos(\epsilon)}{1 - \cos(\epsilon)} \right)$$

where d_{\max} and d_{\min} respectively denote the maximum and minimum Euclidean distance between points in V .

4 Instances in convex position

A finite point set V is in *convex position* when every point in V is a vertex of its convex hull. Deĭneko et al. [5] observed that the Euclidean TSP is easy to solve when V is in convex position. In this case, the optimal permutation is any linear ordering of the points which respects the ordering of the points around the convex hull. Such an ordering can be found in time $O(n \log n)$ [4].

In the context of evolutionary algorithms, the natural question arises, if V is in convex position, how easy is it for simple randomized search heuristics? In this case, a tour is intersection-free if and only if it is globally optimal, hence finding an optimal solution is exactly as hard as finding an intersection-free tour. Since an intersection can be (at least temporarily) removed from a tour by an inversion operation (c.f. Lemma 3), we focus in this section on algorithms that use the inversion operation.

4.1 RLS

RLS operates on a single candidate solution by performing a single inversion in each iteration. Since each inversion which removes an intersection results in a permutation whose fitness is improved by the amount bounded in Lemma 8, it is now straightforward to bound the time it takes for RLS to discover a permutation that corresponds to an intersection-free tour.

Theorem 1. Let V be a set of planar points in convex position angle-bounded by ϵ . The expected time for RLS to solve the TSP on V is $O(n^3 \gamma(\epsilon))$ where γ is defined in Definition 6.

Proof. Consider the infinite stochastic process generated by Algorithm 3 $(x^{(1)}, x^{(2)}, \dots)$. It suffices to bound the expectation of the random variable T defined in Equation (2). Since V is in convex position, any intersection-free tour is globally optimal. Thus, in this case $S_\beta = \emptyset$, and Equation (4) can be written as

$$T = \sum_{t=1}^{\infty} \alpha(x^{(t)}).$$

Consider an arbitrary permutation $x \in S_\alpha$. Since $C(x)$ must contain intersections, by Lemma 3, there is an inversion σ_{ij}^I which removes a pair of intersecting edges and replaces them with a pair of non-intersecting edges. Moreover, by Lemma 8, such an inversion results in an improvement of at least

$$2d_{\min} (1 - \cos(\epsilon)) / (\cos(\epsilon)). \quad (8)$$

If the state x is visited by RLS, the probability that this particular inversion is selected uniformly at random is $2/(n(n-1))$. Thus we have the conditions of Lemma 1 with $p = 2/(n(n-1))$ and c equal to the expression in (8) and the claimed bound on the expectation of T follows. \square

Corollary (to Theorem 1). *If V is in convex position and embedded in an $m \times m$ grid with no three points collinear, then RLS solves the TSP on V in expected time $O(n^3 m^5)$.*

Proof. The bound follows immediately from Theorem 1 since V is in convex position and since, by Lemma 9, V is angle-bounded by $\arctan(1/(2(m-2)^2))$, $d_{max} = (m-1)\sqrt{2}$, and $d_{min} = 1$. Appealing to Lemma 10 yields $\cos(\epsilon)/(1 - \cos(\epsilon)) = O(m^4)$. Substituting these terms into bound of Theorem 1 completes the proof. \square

4.2 The $(\mu + \lambda)$ EA

We now consider the optimization time of the $(\mu + \lambda)$ EA using 2-opt mutation defined in Algorithm 1 applied to a set of points in convex position. Obviously, in this case we will find μ and λ terms appearing in the runtime formulas. We will assume that μ and λ are polynomials in both n and k . We will also find that setting $\lambda = \Theta(\mu n^2)$ ensures that a transition from any state $x^{(t)} \in S_\alpha$ a state which improves on the fitness by at least a specified amount occurs with constant probability.

Such a setting has the effect of reducing the number of *generations* spent removing intersections from the best-so-far tours. However, it is important to note that the number of calls to the fitness function must be accordingly increased by a factor of $\Theta(\mu n^2)$ in each generation. Nevertheless, the expected number of generations can be a useful measure when considering parallel evolutionary algorithms. As Jansen, De Jong and Wegener [15] have pointed out, when the fitness evaluation of the offspring can be performed on parallel processors, the number of generations corresponds to the *parallel* optimization time. In such a case, we would observe a quadratic factor improvement in the parallel runtime corresponding to the segment spend in tours with intersections.

The following theorem bounds the number of expected generations the $(\mu + \lambda)$ EA needs to solve the Euclidean TSP on a set of angle-bounded points in convex position.

Theorem 2. *Let V be a set of planar points in convex position angle-bounded by ϵ . The expected time for the $(\mu + \lambda)$ EA using 2-opt mutation to solve the TSP on V is bounded above by $O((\mu/\lambda) \cdot n^3 \gamma(\epsilon) + n \gamma(\epsilon))$ where γ is as defined in Definition 6.*

Proof. The sequence of best-so-far permutations generated by Algorithm 1 is the infinite stochastic process $(x^{(1)}, x^{(2)}, \dots)$ and we seek the expectation of the random variable T defined in Equation (2). Again, in the case of convex position, there are no non-optimal intersection-free tours so that $S_\beta = \emptyset$ and

$$T = \sum_{t=1}^{\infty} \alpha(x^{(t)}).$$

As long as $x^{(t)}$ is non-optimal, $C(x^{(t)})$ must contain at least one pair of intersecting edges. Hence, in generation t , if $x^{(t)}$ is selected for mutation to create one of the λ offspring, then by Lemma 11, 2-opt mutation must improve the best-so-far solution by at least $c = 2d_{min}(1 - \cos(\epsilon)) / (\cos(\epsilon))$ with probability at least $(en(n-1)/2)^{-1}$. The probability that at least one of the λ offspring improves on $x^{(t)}$ by at least this amount is

$$p \geq 1 - \left(1 - \frac{1}{\mu en(n-1)/2}\right)^\lambda.$$

We now make the following case distinction on λ .

Case $\lambda \geq \mu en(n-1)/2$. For this setting of λ , we have $p \geq 1 - e^{-1}$, so an intersection is removed in each generation with constant probability. Invoking Lemma 1, the expected time to find an intersection-free tour is at most $O(n\gamma(\epsilon))$.

Case $\lambda < \mu en(n-1)/2$. Here we have

$$\left(1 - \frac{1}{\mu en(n-1)/2}\right)^\lambda \geq 1 - e^{-\lambda/(\mu en(n-1)/2)} \geq \frac{\lambda}{\mu en(n-1)}.$$

The final inequality comes from the fact that $e^{-x} \leq 1 - x/2$ for $0 \leq x \leq 1$. Thus, in this case we have $p \geq \lambda/(\mu en^2)$. Applying Lemma 1, the expected time to find an intersection-free tour is at most $O((\mu/\lambda) \cdot n^3\gamma(\epsilon))$.

Thus the expected time to solve the instance is bounded by $O(\max\{(\mu/\lambda) \cdot n^3\gamma(\epsilon), n\gamma(\epsilon)\})$ which yields the claimed bound. \square

Analogous to the Corollary to Theorem 1, we have the following.

Corollary (to Theorem 2). *If V is in convex position and embedded in an $m \times m$ grid with no three points collinear, then RLS solves the TSP on V in expected time $O((\mu/\lambda) \cdot n^3m^5 + nm^5)$.*

5 Parameterized runtime analysis

We now turn our attention to TSP instances in which $\mathfrak{H}(V) \neq V$ and express the expected runtime in terms of $n = |V|$, the number of points, and $k = |V \setminus \mathfrak{H}(V)|$, the number of inner points.

5.1 Expected time to find 2-opt local optima

When $\mathfrak{H}(V) \neq V$, RLS does not necessarily converge to the global optimum with probability one since it can become trapped in a local optimum. However, if the number of inner points is sparse (i.e., $O(1)$), we can bound the complexity of finding local optima. In other words, we want to estimate $E(T_{loc})$, defined in Equation (3).

For RLS, a permutation x is *locally optimum* in the 2-opt neighborhood if there does not exist an inversion $y = \sigma_{ij}^I[x]$ such that $f(y) < f(x)$. In this case, RLS cannot make further improvements. Indeed, if no three points are collinear in V , then if x is locally optimal, $C(x)$ must be intersection-free since, if it were not, Lemma 3 guarantees an inversion exists that removes an intersection and, by the quadrangle inequality, the resulting offspring would have improving fitness. It is important to note that the converse is not necessarily true. However, we may show the following.

Lemma 12. *Suppose $C(x)$ is intersection-free. Then for any neighboring inversion $y = \sigma_{ij}^I[x]$ with $f(y) < f(x)$, $C(y)$ is also intersection-free.*

Proof. Suppose for contradiction that $y = \sigma_{ij}^I[x]$ with $f(y) < f(x)$, but $C(y)$ is not intersection-free. Since $C(x)$ was intersection-free, it follows that the pair of edges introduced by the inversion must intersect. By the quadrangle inequality, the total length of these edges must be greater than the edges they replaced, contradicting that $f(y)$ is strictly less than $f(x)$. \square

Theorem 3. *Suppose V is angle bounded by ϵ and that $|V \setminus \mathfrak{H}(V)| = k$. Then the expected time until RLS finds a local optimum in the 2-opt neighborhood is $O(n^3\gamma(\epsilon) + n^{2k}k!)$ where γ is as defined in Definition 6.*

Proof. After RLS finds an intersection-free tour, by Lemma 12, all subsequent tours will also be intersection free. The total number of intersection-free tours hence serves as a bound on the number of possible improving moves after the first intersection-free tour is encountered. By Lemma 5, this bound is $(n - k)^k k!$.

As long as RLS has not yet found a local optimum, by definition there exists an improving inversion; the expected waiting time to find such an inversion is bounded by $O(n^2)$. Thus, after the first time an intersection-free tour is encountered, the expected time until a local optimum is found is bounded by $O(n^{2k}k!)$.

Finally, for any $x^{(t)} \in S_\alpha$, by Lemma 8 there is a pair (i, j) such that

$$f(\sigma_{ij}^I[x^{(t)}]) < f(x^{(t)}) - 2d_{\min} \left(\frac{1 - \cos(\epsilon)}{\cos(\epsilon)} \right).$$

The probability that RLS selects this inversion is $2/(n(n - 1))$. Substituting these values into Lemma 1 completes the proof. \square

If the points in V are quantized in an $m \times m$ grid, we can appeal directly to Lemmas 9 and 10 to substitute the corresponding angle bounds into the bound obtained in Theorem 3. This results in the following corollary.

Corollary (to Theorem 3). *Suppose that V is quantized in an $m \times m$ grid and that $|V \setminus \mathfrak{H}(V)| = k$. Then the expected time until RLS finds a local optimum is $O(n^3m^5) + O(n^{2k}k!)$.*

5.2 The $(\mu + \lambda)$ EA using 2-opt mutation

The $(\mu + \lambda)$ EA does not suffer from convergence to local optima as does RLS since it uses a Poisson mutation strategy and thus has a non-zero probability of generating any tour (this follows from the connectedness of the inversion adjacency, see, e.g., [12]).

We now use the structural analysis in Section 3 to show that when there are few inner points, intersection-free tours are somehow “close” to an optimal solution in the sense that relatively small perturbations by the EA suffice to solve the problem. Again, recall that we assume μ and λ are polynomials in both n and k .

Theorem 4. *Let V be a set of points angle-bounded by ϵ such that $|V \setminus \mathfrak{H}(V)| = k$. The expected time for the $(\mu + \lambda)$ EA using 2-opt mutation to solve the TSP on V is bounded above by $O((\mu/\lambda) \cdot n^3\gamma(\epsilon) + n\gamma(\epsilon) + (\mu/\lambda) \cdot n^{4k}(2k - 1)!)$.*

Proof. We argue by analyzing the Markov chain generated by the $(\mu + \lambda)$ EA using 2-opt mutation. Let $(x^{(1)}, x^{(2)}, \dots)$ denote the sequence of best-so-far states visited by the $(\mu + \lambda)$ EA.

In generation t , if $x^{(t)} \in S_\alpha$, then the probability of generating an offspring that improves the fitness by at least $c = 2d_{\min}(1 - \cos(\epsilon)) / (\cos(\epsilon))$ is identical to that in the proof of Theorem 2. Arguing in the same manner as in the proof of Theorem 2, we have

$$\sum_{t=1}^{\infty} \alpha(x^{(t)}) = O((\mu/\lambda) \cdot n^3\gamma(\epsilon) + n\gamma(\epsilon)).$$

On the other hand, if $x^{(t)} \in S_\beta$, if it is selected for mutation, then, by Lemma 11, 2-opt mutation produces the optimal solution with probability at least $(en^{4k}(2k-1)!)^{-1}$. Hence, the overall probability of generating an optimal permutation when $x^{(t)}$ is the (intersection-free) population-best permutation is at least

$$q \geq 1 - \left(1 - \frac{1}{\mu en^{4k}(2k-1)!}\right)^\lambda \geq \frac{\lambda}{2\mu en^{4k}(2k-1)!}.$$

Since this is the probability that the Markov chain transits from a state $x^{(t)} \in S_\beta$ to the optimal state, substituting the value for q into the claim of Lemma 2, we have

$$\mathbb{E}\left(\sum_{t=1}^{\infty} \beta(x^{(t)})\right) = O((\mu/\lambda) \cdot n^{4k}(2k-1)!).$$

The bound on $\mathbb{E}(T)$ then follows from Equation 4 and linearity of expectation. \square

Again, from Lemmas 9 and 10 we have the following corollary.

Corollary (to Theorem 4). *Let V be a set of points quantized on an $m \times m$ grid such that $|V \setminus \mathfrak{H}(V)| = k$. The expected time for the $(\mu + \lambda)$ EA using 2-opt mutation to solve the TSP on V is $O((\mu/\lambda) \cdot n^3 m^5 + nm^5 + (\mu/\lambda) \cdot n^{4k}(2k-1)!)$.*

5.3 Mixed mutation strategies

Our analysis so far has revealed important insights into the problem structure of the Euclidean TSP with k inner points and the inversion operator. We now take advantage of these insights to design a new evolutionary algorithm with the aim of explicitly reducing the bound on the expected number of iterations until a permutation corresponding to an optimal TSP tour is found.

The Markov chain analysis relies on the inversion operator to construct an intersection-free tour, but then relies on the inversion operator to simulate the jump operator in order to transform an intersection-free tour into an optimal solution. We now introduce a mutation technique called **mixed-mutation** (outlined in Function 4) that performs both inversion and jump operations, each with constant probability. This allows for a bound which is faster by a factor of $O(n^{2k}(2k-1)!/(k-1)!)$.

Similar to Lemma 11, we have the following result for mixed mutation.

Lemma 13. *Let V be a set of planar points in convex position angle-bounded by ϵ . Then,*

- (1) *For any $x \in S_\alpha$, the probability that mixed mutation creates an offspring y with $f(y) < f(x) - 2d_{\min}(1 - \cos(\epsilon)) / (\cos(\epsilon))$ is at least $(en(n-1))^{-1}$.*
- (2) *For any $x \in S_\beta$, the probability that mixed mutation creates an optimal solution is at least $(2en^{2k}(k-1)!)^{-1}$.*

Proof. For (1), the probability that mixed mutation selects inversions is $1/2$ and the rest of the claim follows from the argument of Lemma 11.

For (2), suppose $x \in S_\beta$. Thus, $C(x)$ is intersection-free, and it follows from Lemma 6 that there are at most k jump operations that transform x into an optimal solution and the remainder of the proof is identical to that of Lemma 11. The probability that mixed mutation selects jump operations contributes the leading factor of 2^{-1} . \square

Function 4: mixed-mutation(x)

input : A permutation x

output: A permutation y

```
1  $y \leftarrow x$ ;  
2 draw  $r$  from a uniform distribution on the interval  $[0, 1]$ ;  
3 draw  $s$  from a Poisson distribution with parameter 1;  
4 if  $r < 1/2$  then  
5    $\lfloor$  perform  $s + 1$  random inversion operations on  $y$ ;  
6 else  
7    $\lfloor$  perform  $s + 1$  random jump operations on  $y$ ;  
8 return  $y$ ;
```

Theorem 5. *Let V be a set of points angle-bounded by ϵ such that $|V \setminus \mathfrak{H}(V)| = k$. The expected time for the $(\mu + \lambda)$ EA using mixed mutation to solve the TSP on V is bounded above by $O((\mu/\lambda) \cdot n^3\gamma(\epsilon) + n\gamma(\epsilon) + (\mu/\lambda) \cdot n^{2k}(k-1)!)$.*

Proof. The proof is identical to the proof of Theorem 4, except we substitute the probabilities from Lemma 13 into Lemmas 1 and 2. \square

As before, from Lemmas 9 and 10 we have the following.

Corollary (to Theorem 5). *Let V be a set of points quantized on an $m \times m$ grid such that $|V \setminus \mathfrak{H}(V)| = k$. The expected time for the $(\mu + \lambda)$ EA using mixed mutation to solve the TSP on V is $O((\mu/\lambda) \cdot n^3m^5 + nm^5 + (\mu/\lambda) \cdot n^{2k}(k-1)!)$.*

6 Conclusion

In this paper, we have studied the runtime complexity of evolutionary algorithms on the Euclidean TSP. We have carried out a parameterized analysis that studies the dependence of the hardness of a problem instance on the number of inner points in the instance. Moreover, we have shown that under reasonable geometric constraints (low angle bounds), simple evolutionary algorithms solve the convex TSP in polynomial time.

In the case that an instance contains k inner points, we have shown that randomized local search using 2-opt can find local optima in expected $O(n^3\gamma(\epsilon)) + O(n^{2k}k!)$ iterations where γ is a function of the angle bound ϵ . For example, when the instance is embedded in an $m \times m$ grid, $\gamma(\epsilon) = O(m^5)$.

Similarly, for the $(\mu + \lambda)$ EA, we have bounded the expected number of generations to solve a TSP instance with k inner points as $O((\mu/\lambda) \cdot n^3\gamma(\epsilon) + n\gamma(\epsilon) + (\mu/\lambda) \cdot n^{4k}(2k-1)!)$. Using the analysis, we have also introduced a mixed mutation strategy based on both permutation jumps and 2-opt moves which attains an improved expected runtime bound of $O((\mu/\lambda) \cdot n^3\gamma(\epsilon) + n\gamma(\epsilon) + (\mu/\lambda) \cdot n^{2k}(k-1)!)$. Hence, with this paper, we have shown that the $(\mu + \lambda)$ EA is a randomized XP-algorithm for the inner point parameterization of Deĭneko et al. [5]. It remains an open question whether or not the algorithm is also a randomized fpt-algorithm.

References

- [1] Anne Auger and Benjamin Doerr. *Theory of Randomized Search Heuristics: Foundations and Recent Developments*. World Scientific Publishing Company, 2011.
- [2] Surender Baswana, Somenath Biswas, Benjamin Doerr, Tobias Friedrich, Piyush P. Kurur, and Frank Neumann. Computing single source shortest paths using single-objective fitness functions. In Thomas Jansen, Ivan Garibay, R. Paul Wiegand, and Annie S. Wu, editors, *Proceedings of the Tenth International Workshop on Foundations of Genetic Algorithms (FOGA'09)*, pages 59–66, Orlando, USA, 2009. ACM Press.
- [3] Barun Chandra, Howard Karloff, and Craig Tovey. New results on the old k -opt algorithm for the traveling salesman problem. *SIAM Journal on Computing*, 28(6):1998–2029, 1999.
- [4] Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars. *Computational Geometry: Algorithms and Applications*. Springer, third edition, 2008.
- [5] Vladimir G. Deineko, Michael Hoffman, Yoshio Okamoto, and Gerhard J. Woeginger. The traveling salesman problem with few inner points. *Operations Research Letters*, 34:106–110, 2006.
- [6] Benjamin Doerr, Edda Happ, and Christian Klein. A tight analysis of the $(1+1)$ -EA for the single source shortest path problem. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC'07)*, pages 1890–1895. IEEE Press, 2007.
- [7] Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity*. Springer, 1999.
- [8] Rodney G. Downey and Michael R. Fellows. Parameterized complexity after (almost) 10 years: Review and open questions. In *Proceedings of Combinatorics, Computation and Logic, DMTCS'99 and CATS'99*, volume 21 of *Australian Computer Science Communications*, pages 1–33, Singapore, 1999. Springer-Verlag.
- [9] Stefan Droste, Thomas Jansen, and Ingo Wegener. On the optimization of unimodal functions with the $(1 + 1)$ evolutionary algorithm. In A. E. Eiben, Thomas Bäck, Marc Schoenauer, and Hans-Paul Schwefel, editors, *Proceedings of the Fifth International Conference on Parallel Problem Solving from Nature (PPSN'98)*, volume 1498 of *Lecture Notes in Computer Science*, pages 13–22. Springer, 1998.
- [10] Stefan Droste, Thomas Jansen, and Ingo Wegener. On the analysis of the $(1+1)$ evolutionary algorithm. *Theoretical Computer Science*, 276:51–81, 2002.
- [11] Matthias Englert, Heiko Röglin, and Berthold Vöcking. Worst case and probabilistic analysis of the 2-opt algorithm for the TSP. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1295–1304. Society for Industrial and Applied Mathematics, 2007.
- [12] Mariko Hagita, Yoshiaki Oda, and Katsuhiro Ota. The diameters of some transition graphs constructed from Hamilton cycles. *Graphs and Combinatorics*, 18(1):105–117, 2002.

- [13] Jun He and Xin Yao. Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*, 127(1):57–85, 2001.
- [14] Jun He and Xin Yao. Towards an analytic framework for analysing the computation time of evolutionary algorithms. *Artificial Intelligence*, 145(1-2):59–97, 2003.
- [15] Thomas Jansen, Kenneth A. De Jong, and Ingo Wegener. On the choice of offspring population size in evolutionary algorithms. *Evolutionary Computation*, 13(4):413–440, 2005.
- [16] Timo Kötzing, Frank Neumann, Heiko Röglin, and Carsten Witt. Theoretical analysis of two ACO approaches for the traveling salesman problem. *Swarm Intelligence*, 6(1):1–21, 2012.
- [17] Stefan Kratsch, Per Kristian Lehre, Frank Neumann, and Pietro Simone Oliveto. Fixed parameter evolutionary algorithms and maximum leaf spanning trees: A matter of mutation. In Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph, editors, *Proceedings of the Eleventh Conference on Parallel Problem Solving from Nature*, volume 6238 of *Lecture Notes in Computer Science*, pages 204–213. Springer, 2010.
- [18] Stefan Kratsch and Frank Neumann. Fixed-parameter evolutionary algorithms and the vertex cover problem. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’09)*, pages 293–300. ACM Press, 2009.
- [19] Dániel Marx. Searching the k -change neighborhood for TSP is $W[1]$ -hard. *Operations Research Letters*, 36(1):31 – 36, 2008.
- [20] F. Neumann and I. Wegener. Randomized local search, evolutionary algorithms, and the minimum spanning tree problem. *Theoretical Computer Science*, 378(1):32–40, 2007.
- [21] Frank Neumann. Expected runtimes of a simple evolutionary algorithm for the multi-objective minimum spanning tree problem. *European Journal of Operational Research*, 181(3):1620–1629, 2007.
- [22] Frank Neumann and Ingo Wegener. Minimum spanning trees made easier via multi-objective optimization. *Natural Computing*, 5(3):305–319, 2006.
- [23] Frank Neumann and Carsten Witt. *Bioinspired Computation in Combinatorial Optimization – Algorithms and Their Computational Complexity*. Springer, 2010.
- [24] Louis V. Quintas and Fred Supnick. On some properties of shortest Hamiltonian circuits. *The American Mathematical Monthly*, 72(9):977–980, 1965.
- [25] Jens Scharnow, Karsten Tinnefeld, and Ingo Wegener. The analysis of evolutionary algorithms on sorting and shortest paths problems. *Journal of Mathematical Modelling and Algorithms*, pages 349–366, 2004.
- [26] Andrew M. Sutton, Jareth Day, and Frank Neumann. A parameterized runtime analysis of evolutionary algorithms for MAX-2-SAT. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’12)*. ACM, 2012. To appear.

- [27] Andrew M. Sutton and Frank Neumann. A parameterized runtime analysis of evolutionary algorithms for the Euclidean traveling salesperson problem. In *Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI'12)*. AAAI Press, 2012. To appear.
- [28] Andrew M. Sutton and Frank Neumann. A parameterized runtime analysis of simple evolutionary algorithms for makespan scheduling. In *Proceedings of the Twelfth International Conference on Parallel Problem Solving from Nature (PPSN'12)*. Springer, 2012. To appear.
- [29] Madeleine Theile. Exact solutions to the traveling salesperson problem by a population-based evolutionary algorithm. In Carlos Cotta and Peter Cowling, editors, *Evolutionary Computation in Combinatorial Optimization*, volume 5482 of *Lecture Notes in Computer Science*, pages 145–155. Springer Berlin / Heidelberg, 2009.
- [30] Jan van Leeuwen and Anneke A. Schoone. Untangling a traveling salesman tour in the plane. In *Proceedings of the Seventh International Workshop on Graph-Theoretic Concepts in Computer Science*, 1981.
- [31] C. Witt. Worst-case and average-case approximations by simple randomized search heuristics. In *Proceedings of the Twenty-Second Symposium on Theoretical Aspects of Computer Science (STACS'05)*, volume 3404 of *Lecture Notes on Computer Science*, pages 44–56. Springer, 2005.
- [32] Yang Yu and Zhi-Hua Zhou. A new approach to estimating the expected first hitting time of evolutionary algorithms. In *Proceedings of the Twenty-First Conference on Artificial Intelligence (AAAI'06)*. AAAI Press, 2006.